ESTIMATING BIVARIATE TAIL

Elena DI BERNARDINO\textsuperscript{b} joint work with
Clémentine PRIEUR\textsuperscript{a} and Véronique MAUME-DESCHAMPS\textsuperscript{b}

\textsuperscript{a}LJK, Université Joseph Fourier, Grenoble 1
\textsuperscript{b}Laboratoire SAF, ISFA, Université Lyon 1
Framework

**Goal**: estimating the tail of a bivariate distribution function.

**Idea**: a general extension of the Peaks-Over-Threshold method.

**Tools**:
- A two-dimensional version of the Pickands-Balkema-de Haan Theorem,
- Yuri & Wüthrich’s approach of the tail dependence.

We present real data examples which illustrate our theoretical results.

**Key words**: Extreme Value Theory, Peaks Over Threshold method, Pickands-Balkema-de Haan Theorem, tail dependence.
Contents

1 One-dimensional results
   • The univariate POT method

2 In dimension 2
   • The framework
   • 2D Pickands-Balkema-de Haan Theorem

3 Estimating the tail of bivariate distributions
   • Construction of the bivariate estimator
   • Convergence results and asymptotic properties

4 Application to real data
Contents

1 One-dimensional results
   • The univariate POT method

2 In dimension 2
   • The framework
   • 2D Pickands-Balkema-de Haan Theorem

3 Estimating the tail of bivariate distributions
   • Construction of the bivariate estimator
   • Convergence results and asymptotic properties

4 Application to real data
Main idea of POT: use of the generalized Pareto distribution (1) to approximate the distribution of excesses over thresholds.

\[ V_{k,\sigma}(x) := \begin{cases} 1 - \left(1 - \frac{kx}{\sigma}\right)^{\frac{1}{k}}, & \text{if } k \neq 0, \sigma > 0, \\ 1 - e^{-\frac{x}{\sigma}}, & \text{if } k = 0, \sigma > 0, \end{cases} \tag{1} \]

and \( x \geq 0 \) for \( k \leq 0 \) or \( 0 \leq x < \frac{\sigma}{k} \) for \( k > 0 \).

A distribution is in the domain of attraction of an extreme value distribution if and only if the distribution of excesses over high thresholds is asymptotically generalized Pareto (e.g. Balkema and de Haan, 1974, Pickands, 1975).
Let \( X_1, X_2, \ldots \) be a sequence of i.i.d random variables with unknown distribution function \( F \).

Fix a threshold \( u \). For \( x > u \), decompose \( F \) as

\[
F(x) = \mathbb{P}[X \leq x] = (1 - \mathbb{P}[X \leq u]) F_u(x - u) + \mathbb{P}[X \leq u],
\]

where \( F_u(x) = \mathbb{P}[X \leq x + u \mid X > u] \).

POT estimate for \( x > u \),

\[
\hat{F}^*(x) = (1 - \hat{F}_X(u)) V_{k,\hat{\sigma}}(x - u) + \hat{F}_X(u).
\]

This univariate modeling is well understood, and has been discussed by Davison (1984), Davison and Smith (1990), McNeil (1997, 1999) and other papers of these authors.
Contents

1 One-dimensional results
   - The univariate POT method

2 In dimension 2
   - The framework
   - 2D Pickands-Balkema-de Haan Theorem

3 Estimating the tail of bivariate distributions
   - Construction of the bivariate estimator
   - Convergence results and asymptotic properties

4 Application to real data
Setting:
- \( X, Y \) two real valued r.v. with continuous df \( F_X \) and \( F_Y \),
- the dependence between \( X \) and \( Y \) is described by a continuous and symmetric copula \( C \).

Notation and definitions:
Survival Copula
\[ \forall (u_1, u_2) \in [0, 1]^2, \ C^*(u_1, u_2) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2). \]

Upper-tail dependence copula \( X, Y \sim \mathcal{U}[0, 1] \), with symmetric \( C \), \( u \in [0, 1) / C^*(1 - u, 1 - u) > 0 \). Then, \( \forall (x, y) \in [0, 1]^2 \), one defines

\[
C_{up}^u(x, y) := \mathbb{P}[X \leq \tilde{F}_u^{-1}(x), Y \leq \tilde{F}_u^{-1}(y) | X > u, Y > u]
\]

with \( \tilde{F}_u(x) := \mathbb{P}[X \leq x | X > u, Y > u] = 1 - \frac{C^*(1 - x \vee u, 1 - u)}{C^*(1 - u, 1 - u)}. \)
Theorem (Upper-tail Theorem; Juri and Wüthrich 2003)

Let $C$ be a symmetric copula such that $C^*(1-u,1-u) > 0$, for all $u > 0$. Furthermore, assume that there is a strictly increasing continuous function $g : [0, \infty) \to [0, \infty)$ such that

$$
\lim_{u \to 1} \frac{C^*(x(1-u), 1-u)}{C^*(1-u, 1-u)} = g(x), \quad x \in [0, \infty).
$$

Then, there exists a $\theta > 0$ such that $g(x) = x^\theta g\left(\frac{1}{x}\right)$ for all $x \in (0, \infty)$. Further, for all $(x, y) \in [0, 1]^2$

$$
\lim_{u \to 1} C^\text{up}_u(x, y) = x + y - 1 + G(g^{-1}(1-x), g^{-1}(1-y)) := C^*G(x, y), \quad (2)
$$

with $G(x, y) := y^\theta g\left(\frac{x}{y}\right)$ $\forall (x, y) \in (0, 1]^2$ and $G \equiv 0$ on $[0, 1]^2 \setminus (0, 1]^2$. 
Auxiliary result

Proposition (de Haan 1970)

\( F_X \in MDA(H_k) \) is equivalent to the existence of a positive measurable function \( a(\cdot) \) such that, for \( 1 - k x > 0 \) and \( k \in \mathbb{R} \),

\[
\lim_{u \to x_F} \frac{1 - F_X(u + x a(u))}{1 - F_X(u)} = \begin{cases} 
(1 - k x)^{\frac{1}{k}}, & \text{if } k \neq 0, \\
\exp(-x), & \text{if } k = 0.
\end{cases}
\tag{3}
\]

\[(2) \text{and}(3) \Rightarrow [ \text{a 2D version of the Pickands-Balkema-de Haan Theorem}]\]

- Juri & Wüthrich (2003) for a symmetric \( C \) and if \( F_X = F_Y \),
- Di Bernardino, Maume-Deschamps and Prieur (2010) for a symmetric \( C \) even if \( F_X \neq F_Y \).
One-dimensional results
In dimension 2
Estimating the tail of bivariate distributions
Application to real data

2D version Pickands-Balkema-de Haan Theorem

Theorem (Di Bernardino, Maume-Deschamps and Prieur 2010)

Let $X, Y$ be real-valued random variables with continuous distribution functions $F_X \neq F_Y$, $C$ be a symmetric copula. Assume $F_X \in MDA(H_{k_1})$, $F_Y \in MDA(H_{k_2})$, and $C$ satisfies the assumptions of the Upper-tail Theorem for some $g$. Define

- $u_Y = F_Y^{-1}(F_X(u))$,
- $x_{F_X} := \sup\{x \in \mathbb{R} \mid F_X(x) < 1\}$, $x_{F_Y} := \sup\{y \in \mathbb{R} \mid F_Y(y) < 1\}$,
- $A := \{(x,y) : 0 < x \leq x_{F_X} - u, 0 < y \leq x_{F_Y} - u_Y\}$.

Then

$$\sup_A \left| \mathbb{P}(X - u \leq x, Y - u_Y \leq y \mid X > u, Y > u_Y) - C^* G(1 - g(1 - V_{k_1, a_1(x)}(u)), 1 - g(1 - V_{k_2, a_2(u_Y)}(y))) \right| \xrightarrow{u \to x_{F_X}} 0,$$

where $V_{k_i,a_i}$ are the GPDs and $a_i(\cdot)$ as in (3), $i = 1, 2$. 

Lyon - Lausanne: 29 November 2010
Contents

1. One-dimensional results
   - The univariate POT method

2. In dimension 2
   - The framework
   - 2D Pickands-Balkema-de Haan Theorem

3. Estimating the tail of bivariate distributions
   - Construction of the bivariate estimator
   - Convergence results and asymptotic properties

4. Application to real data
A new bivariate tail estimator

**Context:** $F$ bivariate df with continuous marginals $F_X, F_Y$. $F$ is assumed to have a stable tail dependence function $l$ that is $\forall x, y \geq 0$, the following limit exists

$$\lim_{t \to 0} t^{-1} \mathbb{P}(1 - F_X(X) \leq tx \text{ or } 1 - F_Y(Y) \leq ty) = l(x, y),$$

see Huang (1992). Then define

$$\lim_{t \to 0} t^{-1} \mathbb{P}(1 - F_X(X) \leq tx, 1 - F_Y(Y) \leq ty) = R(x, y).$$

We have $\forall x, y \geq 0$, $R(x, y) = x + y - l(x, y)$.

**Asymptotic dependence** $\lambda = R(1, 1) \neq 0$, where $\lambda$ is the upper tail dependence coefficient.

**Asymptotic independence** $\forall x, y \geq 0$, $l(x, y) = x + y$. It is equivalent to $\lambda = R(1, 1) = 0$. 
Asymptotic dependence: \( \lambda > 0 \)

Upper-Tail Theorem of Juri & Wüthrich (2003) holds with

\[
g(x) = \frac{x + 1 - l(x, 1)}{2 - l(1, 1)} = \frac{R(x, 1)}{R(1, 1)}, \quad G(x, y) = \frac{x + y - l(x, y)}{2 - l(1, 1)} = \frac{R(x, y)}{R(1, 1)}.
\]

Moreover \( \forall x > 0, \ g(x) = x g(1/x) \) that is \( \theta = 1. \)

We estimate \( g(x) \) with the estimator of \( l \) in Einmahl, Krajina, Serger (2008):

\[
\hat{l}_n(x, y) = \frac{1}{k_n} \sum_{i=1}^{n} 1\{R(X_i) > n - k_n x + 1 \text{ or } R(Y_i) > n - k_n y + 1\},
\]

where \( R(X_i) \) is the rank of \( X_i \) among \( (X_1, \ldots, X_n) \), and \( R(Y_i) \) is the rank of \( Y_i \) among \( (Y_1, \ldots, Y_n) \), \( i = 1, \ldots, n \).
Estimating $g$, $G$ and $\theta$

We estimate $g(x)$ by

$$
\hat{g}(x) = \frac{x+1-n(x,1)}{2-n(1,1)}.
$$

We estimate $G(x, y)$ by

$$
\hat{G}(x, y) = \frac{x+y-n(x,y)}{2-n(1,1)}.
$$

Finally, we estimate the unknown parameter $\theta$ by

$$
\hat{\theta}_x = \frac{\log \hat{g}(x) - \log \hat{g}(1/x)}{\log x}.
$$

In practice, $k$ is "optimized" for each value of $x$. 

Lyon - Lausanne: 29 November 2010

ESTIMATING BIVARIATE TAIL
On simulations

100 samples of size $n = 1000$; mean curve (full line), empirical standard deviation (dashed lines).

**Figure:** Estimators for $\theta$, $(k, \hat{\theta}_x)$ (left) $x = 0.07$, Survival Clayton copula with parameter 1 (right) $x = 5$, Logistic copula with parameter 0.5
Asymptotic independence: $\lambda = 0$

$\lambda = R(1, 1) = 0 \Rightarrow C(u, u) = 1 - 2(1 - u) + o(1 - u), \text{ for } u \to 1$

**Proposition**

*If $\lambda = 0$ and $C$ is a twice continuously differentiable symmetric copula with the determinant of the Hessian matrix of $C$ at $(1, 1)$ different to zero, then*

$$\lim_{u \to 1} \frac{C^\ast(x(1 - u), 1 - u)}{C^\ast(1 - u, 1 - u)} = \frac{b/2(x^2 + 1) + cx}{b + c} \quad \forall x \in [0, \infty),$$

*and $\theta = 2$, with $b = \frac{\partial^2 C}{\partial u^2}(1, 1)$, $c = \frac{\partial C}{\partial u \partial v}(1, 1)$.*

Remark: The assumptions of Proposition are satisfied, for example, in the case of Ali Mikhail-Haq, Frank, Clayton with $a \geq 0$, Independent and Fairlie-Gumbel-Morgenstern copulas.

Remark: one may need going further in the asymptotic development.
On simulations

100 samples of size $n = 1000$; mean curve (full line), empirical standard deviation (dashed lines).

Figure: Estimators for $\theta$, $(k, \hat{\theta}_x)$ (left) $x = 0.8$, Independent copula (right) $x = 0.7$, Clayton copula with parameter 0.05.
Estimating bivariate tail in literature

From Ledford and Tawn (1996):

\[
\hat{F}^*(y_1, y_2) = \exp\{-\hat{l}(\ln(\hat{F}_{Y_1}^*(y_1)), \ln(\hat{F}_{Y_2}^*(y_2)))\},
\]

for high values of \(y_1\) and \(y_2\), where, for instance, \(\hat{F}_{Y_1}^*(y_1)\) (resp. \(\hat{F}_{Y_2}^*(y_2)\)) comes from the univariate POT method.

Approach based on the univariate dependence function of Pickands, 1981; e.g. see Capéraà and Fougères (2000).

In the case \(\lambda = 0\) these methods produce a significant bias. Basically this happens because a bivariate extreme value (G) type dependence structure is assumed to hold in the joint tail of \(F\) above the marginal thresholds.

In order to overcome this problem, Ledford and Tawn (1996, 1997, 1998) introduced this model:

\[
P[Z_1 > z, Z_2 > z] \sim L(z)P[Z_1 > z]^{\frac{1}{\eta}},
\]

where \(L\) is a slowly varying function at infinity, \(\eta \in (0, 1]\), \((Z_1, Z_2)\) with unit Fréchet marginals.
Contrary to Ledford and Tawn’s method, we will propose a model based on regularity conditions of the *Copula* and on the explicit description of the dependence structure in the joint tail (using $g$, $G$ and $\theta$).

Our estimator covers situations less restrictive than dependence or perfect independence above the thresholds (in a different way from the Ledford and Tawn’s method).

Our method is free from the pre-treatment of data because we can work directly with the original general samples without the transformation in Fréchet marginal distributions.
Let $x > u_X$, $y > u_Y$.

For $(s, t)$ in lateral regions $]-\infty, x] \times ]-\infty, u_Y]$ and $] -\infty, u_x] \times ] -\infty, y]$ we approximate the tail by

$$F^*(s, t) = \exp \{-l(-\log F_X(s), -\log F_Y(t))\}.$$

where $l$ is the stable tail dependence function.

In the joint tail $[u_X, x] \times ]u_Y, y]$ we use Juri & Wüthrich’s non parametric modeling (it also works with asymptotic independence $\lambda = 0$).
For a threshold $u$ define $\hat{u}_Y = \hat{F}_Y^{-1}(\hat{F}_X(u))$. Then, for $\hat{k}_X, \hat{\sigma}_X$ (resp. $\hat{k}_Y, \hat{\sigma}_Y$) the MLE based on the excesses of $X$ (resp. $Y$), we estimate $F(x, y)$ by

$$\hat{F}^*(x, y) = A_n (B_n + C_n) + \hat{F}^*_1(u, y) + \hat{F}^*_2(x, \hat{u}_Y) - \frac{1}{n} \sum_{i=1}^n 1\{x_i \leq u, y_i \leq \hat{u}_Y\}$$

with

- $A_n = \frac{1}{n} \sum_{i=1}^n 1\{x_i > u, y_i > \hat{u}_Y\}$,
- $B_n = 1 - \hat{g}_n(1 - V_{k_X, \hat{\sigma}_X}(x - u)) - \hat{g}_n(1 - V_{k_Y, \hat{\sigma}_Y}(y - \hat{u}_Y))$,
- $C_n = \hat{G}_n(1 - V_{k_X, \hat{\sigma}_X}(x - u), 1 - V_{k_Y, \hat{\sigma}_Y}(y - \hat{u}_Y))$,
- $\hat{F}^*_1(u, y) = \exp\{-\tilde{l}_n(-\log(\hat{F}_X(u)), -\log(\hat{F}_Y^*(y)))\}$,
- $\hat{F}^*_2(x, \hat{u}_Y) = \exp\{-\tilde{l}_n(-\log(\hat{F}_X^*(x)), -\log(\hat{F}_Y(\hat{u}_Y)))\}$. 
New tail estimator: construction 2/2

For a threshold $u$ define $\hat{u}_Y = \hat{F}_Y^{-1}(\hat{F}_X(u))$.

Then, for $\hat{k}_X$, $\hat{\sigma}_X$ (resp. $\hat{k}_Y$, $\hat{\sigma}_Y$) the MLE based on the excesses of $X$ (resp. $Y$), we estimate $F(x, y)$ by

$$\hat{F}^*(x, y) = A_n (B_n + C_n) + \hat{F}_1^*(u, y) + \hat{F}_2^*(x, \hat{u}_Y) - \frac{1}{n} \sum_{i=1}^{n} 1\{X_i \leq u, Y_i \leq \hat{u}_Y\}$$

with

- $A_n = \frac{1}{n} \sum_{i=1}^{n} 1\{X_i > u, Y_i > \hat{u}_Y\}$,
- $B_n = 1 - \hat{g}_n(1 - V_{\hat{k}_X, \hat{\sigma}_X}(x - u)) - \hat{g}_n(1 - V_{\hat{k}_Y, \hat{\sigma}_Y}(y - \hat{u}_Y))$,
- $C_n = \hat{G}_n(1 - V_{\hat{k}_X, \hat{\sigma}_X}(x - u), 1 - V_{\hat{k}_Y, \hat{\sigma}_Y}(y - \hat{u}_Y))$,
- $\hat{F}_1^*(u, y) = \exp\{-\hat{l}_n(-\log(\hat{F}_X(u)), -\log(\hat{F}_Y^*(y)))\}$,
- $\hat{F}_2^*(x, \hat{u}_Y) = \exp\{-\hat{l}_n(-\log(\hat{F}_X^*(x)), -\log(\hat{F}_Y(\hat{u}_Y)))\}$. 
For a threshold $u$ define $\hat{u}_Y = F_Y^{-1}(\hat{F}_X(u))$.

Then, for $\hat{k}_X$, $\hat{\sigma}_X$ (resp. $\hat{k}_Y$, $\hat{\sigma}_Y$) the MLE based on the excesses of $X$ (resp. $Y$), we estimate $F(x, y)$ by

$$\hat{F}^*(x, y) = A_n \left( B_n + C_n \right) + \hat{F}_1^*(u, y) + \hat{F}_2^*(x, \hat{u}_Y) - \frac{1}{n} \sum_{i=1}^{n} 1\{X_i \leq u, Y_i \leq \hat{u}_Y\}$$

with

- $A_n = \frac{1}{n} \sum_{i=1}^{n} 1\{X_i > u, Y_i > \hat{u}_Y\}$,
- $B_n = 1 - \hat{g}_n(1 - V_{\hat{k}_X, \hat{\sigma}_X}(x - u)) - \hat{g}_n(1 - V_{\hat{k}_Y, \hat{\sigma}_Y}(y - \hat{u}_Y))$,
- $C_n = \hat{G}_n(1 - V_{\hat{k}_X, \hat{\sigma}_X}(x - u), 1 - V_{\hat{k}_Y, \hat{\sigma}_Y}(y - \hat{u}_Y))$,
- $\hat{F}_1^*(u, y) = \exp\{-\tilde{l}_n(-\log(\hat{F}_X(u)), -\log(\hat{F}_Y^*(y)))\}$,
- $\hat{F}_2^*(x, \hat{u}_Y) = \exp\{-\tilde{l}_n(-\log(\hat{F}_X^*(x)), -\log(\hat{F}_Y(\hat{u}_Y)))\}$. 
Assumptions on the marginals

The assumptions below are assumed both for \( F_X \) and \( F_Y \).

**First order assumptions** \( F \) is in the maximum domain of attraction of Fréchet, that is \( \exists \alpha > 0 \) such that \( \bar{F}(x) = x^{-\alpha}L(x) \) with \( L \) a *slowly varying* function.

**Second order assumptions** as in Smith (1987), we assume that \( L \) satisfies

\[
\text{SR2: } \frac{L(tx)}{L(x)} = 1 + k(t)\phi(x) + o(\phi(x)), \quad \forall \ t > 0, \text{ as } x \to \infty
\]

with \( \phi \) positive and \( \phi(x) \xrightarrow{x \to +\infty} 0 \).
Assumptions on the marginals

The assumptions below are assumed both for $F_X$ and $F_Y$.

**First order assumptions** $F$ is in the maximum domain of attraction of Fréchet, that is $\exists \alpha > 0$ such that $F(x) = x^{-\alpha} L(x)$ with $L$ a *slowly varying* function.

**Second order assumptions** as in Smith (1987), we assume that $L$ satisfies

$$SR2: \, \frac{L(tx)}{L(x)} = 1 + k(t)\phi(x) + o(\phi(x)), \quad \forall \, t > 0, \text{ as } x \to \infty$$

with $\phi$ positive and $\phi(x) \xrightarrow{x \to +\infty} 0$. 
Convergence results

In the asymptotic dependence setting:

For convenient $x_n, y_n$, we prove a Gaussian approximation result for the absolute error (with convergence rate)

$$F^*(x_n, y_n) - \hat{F}^*(x_n, y_n)$$

In the asymptotic independent setting:

$$\lim_{u \to 1} \frac{C^*(x(1-u), 1-u)}{C^*(1-u, 1-u)} = g(x), \ x \in [0, \infty).$$

refined by a second order assumption on copula $C$ (Draisma et al. 2004):

$$\lim_{t \to 0} \frac{\frac{C^*(tx, ty)}{C^*(t,t)} - G(x, y)}{q_1(t)} := Q(x, y)$$

with $x, y \geq 0$, $x + y > 0$, where $q_1$ is some positive function and $Q$ is neither a constant nor a multiple of $G$. 
Contents

1 One-dimensional results
   • The univariate POT method

2 In dimension 2
   • The framework
   • 2D Pickands-Balkema-de Haan Theorem

3 Estimating the tail of bivariate distributions
   • Construction of the bivariate estimator
   • Convergence results and asymptotic properties

4 Application to real data
Loss ALAE and Storm insurance

Real cases to illustrate the finite sample properties of our estimator. In order to study these data we will follow some classical models which make use of a symmetric copula structure.

Figure: Logarithmic scale (left) ALAE versus Loss (right) Storm damages.
We estimate $\theta$ and $F(2.10^5, 10^5)$.
References: Frees and Valdez (1998); Beirlant et al. (2010).

**Figure:** (left) $\hat{\theta}_{0.04}$ (right) $\hat{F}^*(2.10^5, 10^5)$ (full line), $\hat{F}^*(2.10^5, 10^5)$ (dashed line), with the empirical probability indicated with a horizontal line.
Storm insurance


Figure: (left) $\hat{\theta}_{0.05}$ (right) $\hat{F}^*(8.10^3, 950)$ (full line), $\hat{F}^*(8.10^3, 950)$ (dashed line), with the empirical probability indicated with a horizontal line.
Wave surge data

Wave surge data comprising 2894 bivariate events (1971-1977) in Cornwall (England). (Symmetric Copula ?)

Figure: Wave Height (m) versus Surge (m).
Wawe surge data

We estimate $\theta$ and $F(8.32, 0.51)$.

Figure: (left) $\hat{\theta}_{0.02}$ (right) $\hat{F}^*(8.32, 0.51)$ (full line), $\hat{F}^*(8.32, 0.51)$ (dashed line), with the empirical probability indicated with a horizontal line.
Summary

- a new and different approach for estimating bivariate tails,
- we need neither Ledford & Tawn assumptions nor unit Fréchet margins,
- as for L & T estimate, it also works with asymptotic independence.
Ideas for future developments

- use the bivariate tail estimator $\hat{F}^*(x, y)$ to obtain estimation of bivariate upper-quantile curves, for high levels $\alpha$.

- application to the estimation of bivariate Value-at-Risk for large $\alpha$:

$$\text{VaR}_\alpha(\hat{F}) := \{(x, y) \in (\hat{f}_1(n), +\infty) \times (\hat{f}_2(n), +\infty) : \hat{F}^*(x, y) = \alpha\}.$$
Thank for your attention.


